THE JOLLY-SEBER MODEL AND MAXIMUM LIKELIHOOD ESTIMATION REVISITED

BU-847-M by August 1984

Walter K. Kremers

Biometrics Unit, Cornell, Ithaca, New York

ABSTRACT

Seber (1965, *Biometrika* **52**: 249-259) describes a model for a tag-recapture study allowing birth, death, and migration, and derives the likelihood for the model. He then derives supposed maximum likelihood estimates (MLE's) for the parameters of the model. Here we show that in general his estimates are not MLE's. Conditions are given for his estimates to be MLE's and we argue for the use of his estimates over the true MLE's.

1. INTRODUCTION

The ability to estimate population parameters has advanced greatly since the classic papers of Jolly (1965) and Seber (1965). Not only did these papers propose a useful stochastic model for biologists in the estimation of population parameters but they also introduced a general methodology for model building for estimation of population parameters. The estimates given by Seber and by many works to follow, such as those of Buckland (1980) and Pollock (1981), are based upon the well-founded theory of maximum likelihood estimates (MLE's). One difficulty of the estimates in the above works though is that they are not MLE's. We show this explicitly for Seber's estimates and similar arguements apply to the estimates of Buckland and Pollock. The discussion of Seber's model and estimation applies in general to tag-recapture statistics when the objective is to estimate population size, with capture and survival probabilities unknown. We also discuss why we might not want to use MLE's for population size when capture and survival probabilities are unknown. We offer justification for the estimates proposed by Seber, and describe a modification with which Seber's estimates of survival and capture probabilities can be derived through the likelihood approach.

2. THE JOLLY-SEBER MODEL AND ESTIMATION

The Jolly-Seber model allows for a total of s samples or capture periods. For sample i, i=1,...,s, each animal of the sample population is assumed to be captured independently of all other animals with a probability of p_i, after which the animal is immediately identified and returned to the population. Each animal alive immediately after sample i is assumed to survive until the next sample with the probability ϕ_i, independently of all other animals. With notation the same as Seber's let,

a_i = number of animals captured in sample i,

$a_{<i}$ = number of different animals captured before sample i,

$a_{>i}$ = number of different animals captured after sample i,

$a_{<i\cdot i} = a_{<i} + a_i - a_{<i+1}$

$a_{>i\cdot i} = a_{>i} + a_i - a_{>i-1}$

$b_{i+1} = a_{>i} + a_{<i+2} - a_{<s} - a_s - a_{i+1}$

m_1 = population size at the time of sample 1,

$m_i - m_{i-1}$ = net increase in the number of unmarked animals between samples i-1 and i, for i=2,...,s,

$q_i = 1-p_i$

$\alpha_i = \phi_i q_{i+1}$

$\beta_i = \phi_i p_{i+1}$

from which we find the likelihood, $L = L[\{m_i\},p_1,\{\alpha_i,\beta_i\}]$, is proportional to

$$\prod^s \{[(m_i-a_{<i})!/(m_i-a_{<i+1})!]\ p_i^{a_{<i+1}-a_{<i}}\ q_i^{m_i-a_{<i+1}}\}$$
$$x\prod^{s-1}\{\chi_i^{a_i-a_{>i\cdot i}}\ \beta_i^{a_{<i+1\cdot i+1}}\ \alpha_i^{b_{i+1}}\}\ .$$

In general let $\hat{\theta}$ denote the MLE of the parameter θ. Seber equates the difference equations of the form $L[m_i] - L[m_i-1]$ with zero and concludes that,

$$\hat{m}_i = a_{<i+1} + (a_{<i+1} - a_{<i})\hat{\alpha}_{i-1}/\hat{\beta}_{i-1}.$$

However all that may be safely concluded is that $\hat{m}_i$ is less than the quantity to the right of the equals sign and $\hat{m}_i+1$ is greater than the quantity to the right. Using the difference equation improperly Seber claims that the partial derivatives with respect to α_i and β_i respectively are

$$b_{i+1} + \sum^i \{(a_j-a_{>j\cdot j})\hat{\alpha}_j\ldots\hat{\alpha}_i(\hat{\chi}_{i+1}-1)/\chi_j\}=0 \text{ for } i=1,\ldots,s-1 \text{ and,} \quad (1)$$

$$a_{<i+1\cdot i+1} - \sum^i \{(a_j-a_{>j\cdot j})\hat{\alpha}_j\ldots\hat{\alpha}_{i-1}\hat{\beta}_i/\chi_j\}=0 \quad i=1,\ldots,s-1. \quad (2)$$

However most properly

$$\delta L/\delta\alpha_i = \{m_i - a_{<i+1} - (a_{<i+1} - a_{<i})\hat{\alpha}_{i-1}/\hat{\beta}_{i-1}\}/\{(\hat{\alpha}_{i-1}+\hat{\beta}_{i-1})\hat{\alpha}_{i-1}/\hat{\beta}_{i-1}\}$$
$$+ b_{i+1} + \sum^{i} \{(a_j - a_{>j\cdot j})\hat{\alpha}_j \ldots \hat{\alpha}_i(\hat{\chi}_{i+1}^{-1})/\chi_j\}$$

for $i=1,\ldots,s-2$, and

$$\delta L/\delta\beta_i = \{-m_i + a_{<i+1} + (a_{<i+1} - a_{<i})\hat{\alpha}_{i-1}/\beta_{i-1}\}/(\hat{\alpha}_{i-1}+\beta_{i-1})$$
$$+ a_{<i+1\cdot i+1} - \sum^{i} \{(a_j - a_{>j\cdot j})\hat{\alpha}_j \ldots \hat{\alpha}_{i-1}\hat{\beta}_i/\chi_j\}$$

for $i=1,\ldots,s-1$.

Since the first term in $\delta L/\delta\alpha_i$ is nonpositive, if we use Seber's solutions to the maximum-likelihood equations, we find that $\delta L/\delta\alpha_i \leq 0$. Similarly since the first term of $\delta L/\delta\beta_i$ is nonnegative, $\delta L/\delta\beta_i \geq 0$ when evaluated with his solutions to the maximum-likelihood equations.

From this we conclude that the estimates given by Seber for the α_i and β_i are not MLE's and hence neither are his estimates of the m_i. Not only are Seber's estimates not MLE's in the technical sense, but neither are his estimates of the m_i assured of being within one of the true MLE's. Another claim of Seber is that m_1 and p_1 are not separately estimable. The likelihood is however maximized in a trivial way if $m_1=a_1$ and $p_1=1$. In a similar way closed form MLE's for m_s, α_{s-1}, and β_{s-1} are $(a_{<s} + a_s)$, 0 and $a_{>s-1\cdot s-1}/a_{s-1}$ respectively.

3. SEBER'S ESTIMATES AS MLE's

We now consider models for which Seber's estimates are MLE's. If in sample i, each nontagged animal is captured with the probability ρ_i, and each tagged animal is captured with probability p_i, with ρ_i and p_i unrelated, the likelihood is proportional to

$$\Pi^{s} \{[(m_i-a_{<i})!/(m_i-a_{<i+1})!]\ \rho_i^{a_{<i+1}-a_{<i}}\ (1-\rho_i)^{m_i-a_{<i+1}}\}$$
$$x\Pi^{s-1}\{\chi_i^{a_i-a_{>i\cdot i}}\ \beta_i^{a_{<i+1\cdot i+1}}\ \alpha_i^{b_{i+1}}\}\ .$$

For this model $\delta L/\delta\alpha_i$ is given by (1), but for i=1,...,s-2, and $\delta L/\delta\beta_i$ is given by (2) and Seber's estimates of α_i and β_i become MLE's. Note that for this model, in a trivial way, $\hat{m}_i=a_{<i+1}$ and $\hat{p}_i=1$, and hence this model is of little value when estimating the m_i. Seber's estimates are also conditional MLE's when the likelihood is conditioned on the number of animals released into the population as tagged animals. That is, Seber's estimates can be derived from a likelihood approach if we consider the estimation of survival rates estimated solely from the capture of tagged animals

4. MAXIMUM LIKELIHOOD ESTIMATION OF POPULATION SIZE

Seber's estimates of m_i are often found to have a negative bias (Gilbert,1973). The partial derivatives of the likelihood of Seber's model, evaluated at his estimates, satisfy the inequalities $\delta L/\delta\alpha_i \leq 0$ and $\delta L/\delta\beta_i \geq 0$. Hence the MLE's of the m_i are less than his estimates and may incur an even greater bias. The cause of this negative bias of the MLE's when compared to Seber's estimates is the same as that which forces $\hat{m}_1=a_1$ and $\hat{p}_i=1$. The estimation of the p_i and ϕ_i, or of the α_i and β_i, is not independent on the m_i or the $\hat{m}_i$, and the first product of the likelihood cannot be ignored when finding MLE's for the α_i and β_i. The first product of the likelihood is maximized, with respect to m_i and p_i, by $m_i=a_{<i+1}$ and $p_i=1$. Further, given any starting value for p_i and the maximizing m_i, the value of this product can be increased by taking a smaller value for m_i and a larger value for p_i (unless $a_{<i+1}-a_{<i}=0$). Hence the influence of the first product in the estimation of p_i is to pull $\hat{p}_i$ toward one and to

decrease $\hat{m}_i$. This is observed in the most extreme way for m_1, p_1 and m_s, p_s as described above and continues to be undesirable in the estimation of m_i for $i=2,\ldots,s-2$.

For the purpose of discussion we will identify "N" as the population size. In Seber's model an estimate of the population size may be derived from $\hat{m}_i$ along with the $\hat{\alpha}_i$ and $\hat{\beta}_i$. MLE's of N are often justified by an asymptotic argument. However when estimating N, with p and ϕ unknown, the usual asymptotics do not apply. Should we consider the estimation of N as the number of animals sampled becomes stochastically large, (the number of animals caught is a random variable,) if N is not allowed to vary then p must be allowed to go to one. However in practice p may be far from one. Hence to bound p from above by a constant less than one, and to allow number of animals sampled to stochastically increase, is to vary N. Again the parameters being estimated are required to vary in the asymptotic argument, and the usual theory does not apply when we wish to estimate the m_i, α_i and the β_i.

5. PROPOSED METHODOLOGY

Even though Seber's estimates are not MLE's they have a strong intuitive and rational appeal as discussed by Cormack (1972). Also the estimates of the α_i and β_i are the MLE's when estimation is based on the recapture of previously tagged animals, the proper subpopulation from which we should make inferences concerning survival and catchability. If the total number of animals which enter the population of concern is interpreted as the sample size, then we find that Seber's estimates of the α_i and β_i possess desirable asymptotic properties of MLE's, as the $m_i - m_{i-1}$ increase without bound. Let $\tilde{m}_i$ denote Seber's estimates of m_i.

Asymptotics such as those found in Carothers (1973) imply that the relative biases of the $\tilde{m}_i$ approach zero as the $m_i - m_{i-1}$ are allowed to increase without bound (when the assumptions of the model are met). This, along with the asymptotic variances found in Seber, implies the $[\ E(\tilde{m}_i - m_i)^2\ ]^{1/2} / m_i$ approach zero as the $m_i - m_{i-1}$ are allowed to increase without bound, suggesting a type of weak consistency.

6. CONCLUSION

The failure in Seber's derivation of the MLE's stems from the use of a difference equation as a differential equation, in particular the equating of $\hat{m}_i - a_{<i+1} - (a_{<i+1} - a_{<i})\hat{\alpha}_{i-1}/\hat{\beta}_{i-1}$ with zero in the evaluation of $\delta L/\delta\alpha_i$ and $\delta L/\delta\beta_i$, rather than considering the proper inequalities. This problem does not arise when population size is the only unknown parameter. The population size is then estimated by the integer part of the solution to the difference equation, thus estimating population size by a true member of the parameter space. However little concern is usually voiced if population size is estimated by the solution to the difference equation itself as this estimate is within one of the MLE. When the likelihood involves unknown parameters in addition to those determining population size, however, one can no longer expect to be within one of the MLE's when the difference equation is set to zero.

Even though the estimates proposed by Seber are not MLE's they do appear to be the proper estimates to use. His estimates of survival and capture probabilities are those we obtain if we consider MLE's for the subpopulation of tagged animals. Further the MLE's of the m_i are biased negatively to a larger extent than Seber's estimates are often found to be through asymptotic expansion and simulation.

7. REFERENCES

Buckland, S.T. (1980). A modified analysis of the Jolly-Seber capture-recapture model. *Biometrics* **36**: 419-435.

Carothers, A.D. (1973). The effects of unequal catchability on Jolly-Seber estimates. *Biometrics* **29**: 79-100.

Cormack, R.M. (1972). The logic of capture-recapture estimates. *Biometrics* **28**: 337-343.

Gilbert,R.O. (1973). Approximations of the bias in the Jolly-Seber capture-recapture model. *Biometrics* **29**: 501-526.

Jolly, G.M. (1965). Explicit estimates from capture-recapture data with both death and immigration--stochastic model. *Biometrika* **52**: 225-247.

Pollock, K.H. (1981). Capture-recapture models allowing for age-dependent survival and capture rates. *Biometrics* **37**: 521-529.

Seber, G.A.F. (1965). A note on the multiple recapture census. *Biometrika* **52**: 249-259.